

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2001, 04:20:59 ; Search time 1474.31 Seconds  
(without alignments)  
138.851 Million cell updates/sec

Title: PCT-US00-32259-12  
Perfect score: 40  
Sequence: 1 ctgagccaccataacccataactccaggattggg 40

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues  
Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_bal:\*
- 2: gb\_ba2:\*
- 3: gb\_om:\*
- 4: gb\_ov:\*
- 5: gb\_ph:\*
- 6: gb\_pl1:\*
- 7: gb\_pl2:\*
- 8: gb\_pr1:\*
- 9: gb\_pr2:\*
- 10: gb\_pr3:\*
- 11: gb\_ro:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: em\_fun:\*
- 15: em\_hum1:\*
- 16: em\_hum2:\*
- 17: em\_in:\*
- 18: em\_om:\*
- 19: em\_or:\*
- 20: em\_ov:\*
- 21: em\_pat:\*
- 22: em\_ph:\*
- 23: em\_pl:\*
- 24: em\_ro:\*
- 25: em\_sts:\*
- 26: em\_sy:\*
- 27: em\_un:\*
- 28: em\_vi:\*
- 29: gb\_htg1:\*
- 30: gb\_htg2:\*
- 31: gb\_in1:\*
- 32: gb\_in2:\*
- 33: em\_ba1:\*
- 34: em\_ba2:\*
- 35: em\_hum3:\*
- 36: em\_hum4:\*
- 37: gb\_pr4:\*
- 38: gb\_htg3:\*
- 39: gb\_htg4:\*
- 40: gb\_htg5:\*
- 41: gb\_htg6:\*
- 42: gb\_htg7:\*
- 43: em\_htg1:\*

- 44: em\_htg2:\*
- 45: em\_htg3:\*
- 46: em\_hum5:\*
- 47: gb\_pl3:\*
- 48: gb\_pr5:\*
- 49: gb\_htg8:\*
- 50: gb\_htg9:\*
- 51: gb\_htg10:\*
- 52: gb\_htg11:\*
- 53: gb\_htg12:\*
- 54: gb\_htg13:\*
- 55: gb\_htg14:\*
- 56: gb\_in3:\*
- 57: gb\_htg15:\*
- 58: gb\_htg16:\*
- 59: gb\_htg17:\*
- 60: em\_htg4:\*
- 61: em\_htg5:\*
- 62: em\_htg6:\*
- 63: em\_htg7:\*
- 64: em\_hum6:\*
- 65: gb\_htg18:\*
- 66: gb\_htg19:\*
- 67: gb\_htg20:\*
- 68: gb\_htg21:\*
- 69: gb\_htg22:\*
- 70: gb\_htg23:\*
- 71: gb\_vil:\*
- 72: gb\_vil2:\*
- 73: gb\_ba3:\*
- 74: em\_htg8:\*
- 75: em\_htg9:\*
- 76: em\_htg10:\*
- 77: gb\_pr6:\*
- 78: gb\_pr7:\*
- 79: gb\_sts1:\*
- 80: gb\_sts2:\*
- 81: gb\_pat1:\*
- 82: gb\_pat2:\*
- 83: em\_htg0:\*
- 84: gb\_htg24:\*
- 85: gb\_pr8:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match %	Length	DB ID	Description
c	1	23	57.5	1653	37	AK026605 Homo sapi
c	2	23	57.5	2600	78	X75861 H.sapiens T
c	3	23	57.5	2634	81	AR058920 Sequence
	4	23	57.5	2634	81	AR058921 Sequence
	5	23	57.5	197189	42	AC019168 Homo sapi
	6	22.8	57.0	171185	70	AP002424 Homo sapi
	7	22.8	57.0	177097	70	AP001569 Homo sapi
c	8	22.8	57.0	179726	9	AC007052 Homo sapi
	9	22.8	57.0	188357	37	AL159996 Human DNA
	10	22.8	57.0	200774	70	AP001592 Homo sapi
	11	22.4	56.0	151514	77	HS471M1.3
	12	22	55.0	452	79	G22123 human STS W
c	13	22	55.0	34887	56	TBR012:99
c	14	22	55.0	76727	77	HS821D:1
c	15	22	55.0	110387	84	HSAP96:1
c	16	22	55.0	177540	9	AC006538
c	17	21.8	54.5	154599	41	AC016997 Homo sapi
c	18	21.8	54.3	155907	49	AC020991 Homo sapi
	19	21.6	54.0	37470	48	CNS00YVD
c	20	21.6	54.0	104436	77	HS611N7
	21	21.6	54.0	168925	68	AL356502 Homo sapi

```
c 22 21.6 54.0 208454 69 AL445199 Homo sapi
23 21.6 54.0 225653 69 AL359920 Homo sapi
24 21.2 53.0 70594 40 AC015786 Homo sapi
c 25 21.2 53.0 125219 69 AL445228 Homo sapi
26 21.2 53.0 147009 29 AC007618 Homo sapi
c 27 21.2 53.0 181215 49 AC021300 Homo sapi
28 21.2 53.0 208992 35 AC061709 Homo sapi
c 29 21.2 53.0 233454 9 AC005703 Homo sapi
30 21 52.5 22021 11 AF109155 Mus muscu
c 31 21 52.5 16332 1 AF104912 Escherich
32 21 52.5 80514 8 AC004454 Homo sapi
c 33 21 52.5 87235 50 AC022963 Homo sapi
34 21 52.5 92215 41 AC018277 Homo sapi
35 21 52.5 111269 30 AC009251 Homo sapi
36 21 52.5 127409 37 AL137790 Human DNA
37 21 52.5 132571 69 AL445245 Homo sapi
38 21 52.5 155278 78 HSDJ655C4 Homo sapi
c 39 21 52.5 167277 52 AC025318 Homo sapi
40 21 52.5 167356 70 AP002814 Homo sapi
41 21 52.5 169480 65 AC0079608 Homo sapi
42 21 52.5 173333 68 AC1355990 Homo sapi
43 21 52.5 181233 38 AC105335 Homo sapi
44 21 52.5 186716 41 AC018468 Homo sapi
c 45 21 52.5 200627 42 AC019225 Homo sapi
```

## ALIGNMENTS

```
RESULT 1
AK026605/c 1653 bp mRNA PRI 29-SEP-2000
LOCUS Homo sapiens cDNA: FLJ22952 fis, clone KAT09742.
DEFINITION AK026605
ACCESSION AK026605
VERSION AK026605.1 GI:10439494
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to mRNA, clone_lib:KAT clone:KAT09742.
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished (2000)
2 (bases 1 to 1653)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo,
Laboratory of Genome Structure Analysis, Human Genome Center;
Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdnal@ms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
```

```
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
```

```
FEATURES             Location/Qualifiers
     source            1..1653
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="KATO III"
                     /cell_type="signet-ring cell carcinoma"
                     /clone="KAT09742"
                     /clone_lib="KAT"
                     /note="cloning vector pME18SFL3"
```

```
BASE COUNT 441 a 369 c 358 g 485 t
ORIGIN
Query Match 57.5%; Score 23; DB 37; Length 1653;
Best Local Similarity 74.4%; Pred. No. 23;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 2 tcgagccaccctcaaccctcaatactccaggattggg 40
||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 869 TCCACCCACACTTGACCCACAAAACTCGAGGGATTGGG 831
```

## RESULT 2

```
HSTEGT/c 2
LOCUS HSTEGT 2600 bp mRNA PRI 28-SEP-1995
DEFINITION H.sapiens TEGT gene.
ACCESSION X75861
VERSION X75861.1 GI:456258
KEYWORDS TEGT gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2600)
Walter,L.
Direct Submission
Submitted (14-JAN-1994) L. Walter, Abteilung Immunogenetik der
Universitaet Goettingen, Gosslerstr 12a, 37073 Goettingen, FRG
2 (bases 1 to 2600)
Walter,L., Marynen,P., Szpirer,J., Levan,G. and Gunther,E.
Identification of a novel conserved human gene, TEGT
Genomics 28 (2), 301-304 (1995)
MEDLINE 96015061
FEATURES             Location/Qualifiers
     source            1..2600
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12q12-q13"
                     /clone_lib="lambda gt11 cDNA"
                     /tissue_type="testis"
                     41..754
                     /gene="TEGT"
                     41..754
                     /gene="TEGT"
                     /codon_start=1
                     /protein_id="CAA53472.1"
                     /db_xref="GI:458545"
                     /db_xref="SWISS-PROT:P35061"
                     /translation="MNIFDKINFDAALLAFSHITPSTQOHLKKVYASFALCMFVAAG
                     AYVWYTHFIQAGLLSALGSLIIMWIMATPHSHETEOKRLGLAGFAFLTGVGLGPA
                     LEFCIAVNPISLPTAFWGTAMIFCTLSALYARRSYLFLGGLMSALSLLISLG
                     NVFFGSINPQANLYGLVYMGFVLVDQLIEKAHEGDQDIWHICIDFLDFTIVF
                     RKLMILIANNEKDKKEKK"
```

## BASE COUNT

```
619 a 621 c 590 g 770 t
ORIGIN
Query Match 57.5%; Score 23; DB 78; Length 2600;
Best Local Similarity 74.4%; Pred. No. 23;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 2 tcgagccaccctcaaccctcaatactccaggattggg 40
||| ||||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2024 TCCACCCACACTTGACCCACAAAACTCGAGGGATTGGG 1986
```

## RESULT 3

```
AR058920/c 3
LOCUS AR058920 2634 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5937838.
ACCESSION AR058920
VERSION AR058920.1 GI:5984497
```





\* NOTE: This is a 'working draft' sequence.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

FEATURES             Location/Qualifiers
     source            1..171185
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="18"
                        /clone="RP11-873L22"
                        /map="18q21"
     misc_feature       1..24923
                        /note="assembly_fragment"
     misc_feature      25024..47004
                        /note="assembly_fragment"
     misc_feature      47105..67070
                        /note="assembly_fragment"
     misc_feature      67171..82757
                        /note="assembly_fragment"
     misc_feature      82858..95544
                        /note="assembly_fragment"
     misc_feature      95645..109592
                        /note="assembly_fragment"
     misc_feature     109693..120167
                        /note="assembly_fragment"
     misc_feature     120268..129255
                        /note="assembly_fragment"
     misc_feature     129356..137003
                        /note="assembly_fragment clone_end:SP6 vector_side:right"
     misc_feature     137104..144376
                        /note="assembly_fragment"
     misc_feature     144477..150659
                        /note="assembly_fragment"
     misc_feature     150760..155896
                        /note="assembly_fragment"
     misc_feature     155997..160522
                        /note="assembly_fragment"
     misc_feature     160623..163263
                        /note="assembly_fragment"
     misc_feature     163364..165194
                        /note="assembly_fragment clone_end:T7 vector_side:left"
     misc_feature     165295..167704
                        /note="assembly_fragment"
     misc_feature     167805..169750
                        /note="assembly_fragment"
     misc_feature     169851..171185
                        /note="assembly_fragment"
BASE COUNT    52031 a 32086 c 33289 g 52078 t 1701 others
ORIGIN
Query Match          57.0%; Score 22.8; DB 70; Length 171185;
Best Local Similarity 79.4%; Pred No. 27;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ctcgagccaccataaccctcaatactccagg 34
    ||| ||| ||||| ||||| ||||| |||||
Db 93406 CTCAGCTCCCGCCATGCCCTCATTTCCAGTG 93439

RESULT 7
LOCUS      AP001569
DEFINITION Homo sapiens chromosome 18 clone RP11-859C31 map 18q21, WORKING
DRAFT SEQUENCE, in unordered pieces.
ACCESSION  AP001569
VERSION     AP001569.2 GI:8117403
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 177097)

```

# AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Homo sapiens 177,097 genomic DNA of 18q21  
 Published Only in DataBase (2000) In press  
 2 (bases 1 to 177097)  
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Direct Submission  
 Submitted (29-MAR-2000) to the DDBJ/EMBL/GenBank databases.  
 Masahira Hattori, The Institute of Physical and Chemical Research  
 (RIKEN), Genomic Sciences Center (GSC); Kitasato Univ., 1-15-1  
 Kitasato, Sagamihara, Kanagawa 228-8555, Japan  
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,  
 Tel:81-42-778-9923, Fax:81-42-778-9924)  
 On May 30, 2000 this sequence version replaced gi:7380904.  
 ----- Genome Center  
 Center: RIKEN Genomic Sciences Center (GSC)  
 Center code: RIKEN  
 Web site: http://hgp.gsc.riken.go.jp/  
 Contact: hattori@gsc.riken.go.jp  
 ----- Project Information  
 Center project name: HumTrafit18  
 Center clone name: RP11-859C21  
 ----- Summary Statistics  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 160670 bases at least Q40  
 Consensus quality: 168058 bases at least Q30  
 Consensus quality: 171631 bases at least Q20  
 Insert size: 173997; sum-of-contigs  
 Quality coverage: 4.63x in Q20 bases; sum-of-contigs

-----

## COMMENT

NOTE: This is a 'working draft' sequence. It currently consists of  
 32 contigs. The true order of the pieces is not known and their  
 order in this sequence record is arbitrary. Gaps between the  
 contigs are represented as runs N, but the exact sizes of the gaps  
 are unknown. This record will be updated with the finished sequence  
 as soon as it is available and the accession number will be  
 preserved.

```

1 23792 contig cf 23792 bp in length
2393 45465 contig cf 21573 bp in length
45566 58167 contig cf 12602 bp in length
58268 71173 contig cf 12906 bp in length
71274 81253 contig cf 9980 bp in length
81354 91080 contig cf 9727 bp in length
91481 97387 contig cf 6207 bp in length
97488 104902 contig cf 7415 bp in length
105003 111107 contig cf 6105 bp in length
111208 116851 contig cf 5644 bp in length
116952 123585 contig cf 6634 bp in length
123686 128985 contig cf 5300 bp in length
129086 134233 contig cf 5148 bp in length
134334 138588 contig cf 4255 bp in length
138689 142573 contig cf 3885 bp in length
142674 146286 contig cf 3613 bp in length
146397 150053 contig cf 3667 bp in length
150154 154046 contig cf 3893 bp in length
154147 156225 contig cf 2079 bp in length
156326 158578 contig cf 2253 bp in length
158679 161124 contig cf 2446 bp in length
161225 163559 contig cf 2335 bp in length
163660 165538 contig cf 1879 bp in length
165639 166949 contig cf 1311 bp in length
167050 168240 contig cf 1191 bp in length
168341 169579 contig cf 1239 bp in length
169680 170978 contig cf 1299 bp in length
171079 172080 contig cf 1002 bp in length
172181 173514 contig cf 1334 bp in length
173615 174614 contig cf 1000 bp in length
174715 175976 contig cf 1262 bp in length
176077 177097 contig cf 1021 bp in length
Sequence updated (26-May-2000).

```

\* NOTE: This is a 'working draft' sequence.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

## FEATURES

```
source
  Location/Qualifiers
    1..177097
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="18"
      /clone="RP11-859C21"
      /map="18q21"
    1..23792
      /note="assembly_fragment"
    23893..45465
      /note="assembly_fragment"
    45566..58167
      /note="assembly_fragment"
    58268..71173
      /note="assembly_fragment"
    71274..81253
      /note="assembly_fragment"
    81354..91080
      /note="assembly_fragment"
    91181..97387
      /note="assembly_fragment"
    97488..104902
      /note="assembly_fragment clone_end:SP6 vector_side:left"
    105003..111107
      /note="assembly_fragment"
    111208..116851
      /note="assembly_fragment"
    116952..123585
      /note="assembly_fragment"
    123686..128985
      /note="assembly_fragment"
    129086..134233
      /note="assembly_fragment"
    134334..138588
      /note="assembly_fragment"
    138689..142573
      /note="assembly_fragment"
    142674..146286
      /note="assembly_fragment clone_end:T7 vector_side:left"
    146387..150053
      /note="assembly_fragment"
    150154..154046
      /note="assembly_fragment"
    154147..156225
      /note="assembly_fragment"
    156326..158578
      /note="assembly_fragment"
    158679..161124
      /note="assembly_fragment"
    161225..163559
      /note="assembly_fragment"
    163660..165538
      /note="assembly_fragment"
    165639..166949
      /note="assembly_fragment"
    167050..168240
      /note="assembly_fragment"
    168341..169579
      /note="assembly_fragment"
    169680..170978
      /note="assembly_fragment"
    171079..172080
      /note="assembly_fragment"
    172181..173514
      /note="assembly_fragment"
    173615..174614
      /note="assembly_fragment"
    174715..175976
      /note="assembly_fragment"
```

```
misc_feature 176077..177097
              /note="assembly_fragment"
BASE COUNT 53531 a 33044 c 34236 g 53184 t 3102 others
ORIGIN
Query Match 57.0%; Score 22.8; DB 70; Length 177097;
Best Local Similarity 79.4%; Pred. No. 27;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 ctcgagccatccataaccctcaatactccagg 34
    ||| |||| ||||| ||||| ||||| |||||
Db 3709 CTCAGGCTCCCATGCCCTCATTTCCAGTG 3742

RESULT 8
AC007052/c
LOCUS AC007052 179726 bp DNA PRI 23-MAR-1999
DEFINITION Homo sapiens chromosome 18, clone hRPK.411_H_24, complete sequence.
ACCESSION AC007052
VERSION AC007052.4 GI:4510438
KEYWORDS HTG.
SOURCE human.
  ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 (bases 1 to 179726)
      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  AUTHORS
    Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE
    Homo sapiens chromosome 18, clone hRPK.411_H_24
  JOURNAL
    Unpublished
  REFERENCE
    2 (bases 1 to 179726)
      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
      Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
      Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
      Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
      Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
      Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
      Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
      Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
      Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
      Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
      Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
      Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
      Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
      Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
      Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  DIRECT SUBMISSION
  TITLE
    Submitted (09-MAR-1999) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
  JOURNAL
    Research, 320 Charles Street, Cambridge, MA 02141, USA
  REFERENCE
    3 (bases 1 to 179726)
      Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
      Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,
      Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
      Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
      Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,
      Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
      Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
      Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P.,
      Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
      Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,
      Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
      Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
      Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
      Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
      Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
  DIRECT SUBMISSION
  TITLE
    Submitted (23-MAR-1999) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
  JOURNAL
    Research, 320 Charles Street, Cambridge, MA 02141, USA
  COMMENT
    On Mar 24, 1999 this sequence version replaced gi:4432872.
    All repeats were identified using RepeatMasker: Smit, A.F.A. &
    Green, E. (1996-1997)
    http://ftp.genome.washington.edu/RM/RepeatMasker.html.
  FEATURES
    Location/Qualifiers
```



```

Best Local Similarity 79.4%; Pred. No. 27;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ctcgagcccccataaacctcaataactcagg 34
||||| ||||| ||||| ||||| ||||| ||||| |
Db 78718 CTCAGCCTCCCATGCCCTCATTTCTCCAGTG 78685

RESULT 9
AL159996 183357 bp DNA PRI 24-JUL-2000
LOCUS Human DNA sequence from clone RP11-401F12 on chromosome 9. Contains
DEFINITION STSS and GSSs, complete sequence.
ACCESSION AL159996
VERSION AL159996.7 GI:8648450
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 183357)
Laird,G.
Direct Submission
Submitted (19-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 21, 2000 this sequence version replaced gi:8346239.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-401F12 is from the library RPCI-11.2 constructed at the
Roswell Park Cancer Institute by the group of Pletier de Jong. For
further details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
This sequence is the entire insert of clone RP11-401F12. The true
left end of clone RP11-208G24 is at 167549 in this sequence. The
true right end of clone RP11-395D3 is at 5200 in this sequence.
Location/Qualifiers
1. .188357
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-401F12"
/clone_lib="RPCI-11.2"
410. .897
/note="match: GSS: Em:AQ606829"
1734. .2348
/note="match: GSS: Em:AQ823967"
1819. .2460
/note="L1MC4 repeat: matches 6462. .7077 of consensus"
2551. .2614
/note="2 copies 32 mer 96% conserved"
2690. .2698
/note="TN10 excised: This sequence represents the
duplicated flanking sequence of the TN10."

repeat_region 3529. .3560
/note="8 copies 4 mer gtgt 96% conserved"
repeat_region 3746. .3959
/note="L1PB3 repeat: matches 5916. .6147 of consensus"
repeat_region 3969. .4006
/note="U2 repeat: matches 1. .38 of consensus"
repeat_region 4392. .4430
/note="13 copies 3 mer aac 92% conserved"
misc_feature 5217. .5472
/note="match: GSS: Em:AQ302480"
misc_feature 5260. .6024
/note="match: GSS: Em:AQ901451"
misc_feature complement(5743. .6091)
/note="match: GSS: Em:AQ132828"
misc_feature 6101. .6658
/note="match: GSS: Em:AQ830958"
repeat_region 7246. .8450
/note="Rigger3b repeat: matches 8. .1231 of consensus"
repeat_region 8626. .8773
/note="L2 repeat: matches 2452. .2611 of consensus"
repeat_region 10391. .10584
/note="MER5A repeat: matches 1. .189 of consensus"
repeat_region 11880. .12174
/note="L2 repeat: matches 1986. .2272 of consensus"
misc_feature complement(12504. .12890)
/note="match: STS: Em:HS311YCI"
repeat_region 12609. .12648
/note="10 copies 4 mer gtgt 90% conserved"
repeat_region 12610. .12647
/note="19 copies 2 mer tg 92% conserved"
repeat_region 12615. .12650
/note="6 copies 6 mer gtgtgt 91% conserved"
repeat_region 13805. .14099
/note="Alusq repeat: matches 2. .295 of consensus"
repeat_region 14101. .14388
/note="Alusq repeat: matches 16. .302 of consensus"
repeat_region 15605. .15839
/note="L2 repeat: matches 2357. .2612 of consensus"
repeat_region 15947. .16013
/note="L2 repeat: matches 2609. .2674 of consensus"
repeat_region 16084. .16364
/note="AluJo repeat: matches 1. .288 of consensus"
misc_feature complement(16254. .16783)
/note="match: GSS: Em:AQ764225"
repeat_region 17436. .17528
/note="MIR repeat: matches 35. .123 of consensus"
repeat_region 18177. .18298
/note="FLAM_A repeat: matches 1. .122 of consensus"
repeat_region 19380. .19675
/note="Alusx repeat: matches 1. .296 of consensus"
misc_feature complement(19534. .19956)
/note="match: GSS: Em:AQ042003"
repeat_region 19736. .19958
/note="L2 repeat: matches 2477. .2745 of consensus"
repeat_region 20126. .20237
/note="L2 repeat: matches 2589. .2705 of consensus"
repeat_region 20737. .20812
/note="WAE1 repeat: matches 5. .80 of consensus"
repeat_region 21025. .21200
/note="L2 repeat: matches 2184. .2355 of consensus"
repeat_region 21443. .21646
/note="MIR repeat: matches 40. .246 of consensus"
repeat_region 23474. .23793
/note="MLTIF repeat: matches 197. .541 of consensus"
misc_feature complement(25531. .26230)
/note="match: GSS: Em:AQ787911"
repeat_region 27777. .28867
/note="L1MC3 repeat: matches 6667. .7734 of consensus"
repeat_region 28802. .28880
/note="L1MC2 repeat: matches 6247. .6326 of consensus"
repeat_region 28924. .29133
/note="MER20 repeat: matches 2. .214 of consensus"
repeat_region 29810. .29849

```



```
/note="20 copies 2 mer ac 80% conserved"
29868..30157
/note="AluY repeat: matches 3..290 of consensus"
30822..31276
/note="MLT1H repeat: matches 70..532 of consensus"
31655..32020
/note="THE1B repeat: matches 1..364 of consensus"
32181..32674
/note="match: GSS: Em:AQ822292"
32279..32403
/note="MIR repeat: matches 57..192 of consensus"
complement(32857..33263)
/note="match: GSS: Em:AQ139531"
33279..33551
/note="AluSg repeat: matches 1..301 of consensus"
35349..35388
/note="10 copies 4 mer acac 92% conserved"
35353..35388
/note="3 copies 12 mer 94% conserved"
35759..36002
/note="MIR repeat: matches 2..246 of consensus"
36104..36469
/note="match: GSS: Em:AQ031409"
37497..37545
/note="MIR repeat: matches 81..127 of consensus"
37684..38248
/note="MLT2D repeat: matches 1..553 of consensus"
38259..38373
/note="MIR repeat: matches 122..249 of consensus"
36334..38392
/note="L2 repeat: matches 2692..2750 of consensus"
38609..38887
/note="match: STS: Em:G13414"
38781..38991
/note="LTR33 repeat: matches 302..518 of consensus"
39086..39212
/note="LTR33 repeat: matches 29..165 of consensus"
39249..39483
/note="MER4D repeat: matches 1..234 of consensus"
39483..40142
/note="MER4D repeat: matches 359..1017 of consensus"
40325..40415
/note="MER94 repeat: matches 39..134 of consensus"
41795..41902
/note="LTR33 repeat: matches 370..472 of consensus"
41903..42192
/note="AluSg repeat: matches 1..291 of consensus"
42193..42241
/note="LTR33 repeat: matches 472..521 of consensus"
43494..43613
/note="L2 repeat: matches 2628..2748 of consensus"
43695..43862
/note="FRAM repeat: matches 0..164 of consensus"
44019..44082
/note="L2 repeat: matches 2356..2428 of consensus"
44083..44384
/note="L1P4 repeat: matches 5841..6144 of consensus"
44385..44529
/note="L2 repeat: matches 2252..2356 of consensus"
44541..44628
/note="MIR repeat: matches 128..229 of consensus"
44631..44747
/note="MER33 repeat: matches 203..324 of consensus"
44779..44985
```

Query Match 57.0%; Score 22.8; DB 37; Length 188357;  
Best Local Similarity 79.4%; Pred. NO. 27;  
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 7 ccacccataaccctcaatactccaggattggg 40  
||||| | | | | | | | | | | | | | | | | |  
Db 56526 CCACCCCAATGTCCTCAAGACTCCAGGGGAGGG 56559

RESULT 10  
AP001592  
LOCUS  
DEFINITION  
Homo sapiens chromosome 18 clone RP11-869L2 map 18q21, WORKING  
DRAFT SEQUENCE, in unordered pieces.  
AP001592  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AP001592 200774 bp DNA HTG 29-JUL-2000  
Homo sapiens chromosome 18 clone RP11-869L2 map 18q21, WORKING  
DRAFT SEQUENCE, in unordered pieces.  
AP001592  
AP001592.3 GI:9588073  
HTG; HTGS PHASE1; HTGS DRAFT.  
Homo sapiens DNA, clone:RP11-869L2.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 200774)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
Homo sapiens 200,774 genomic DNA of 18q21  
Published Only in DataBase (2000) In press  
2 (bases 1 to 200774)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
Direct Submission  
Submitted (31-MAR-2000) to the DDBJ/EMBL/GenBank databases.  
Masahira Hattori, The Institute of Physical and Chemical Research  
(RIKEN), Genomic Sciences Center (GSC); 1-15-1 Kitasato,  
Sagamihara, Kanagawa 228-8555, Japan  
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,  
Tel:81-42-778-9923, Fax:81-42-778-9924)  
On Jul 31, 2000 this sequence version replaced gi:8117416.  
----- Genome Center

Center: RIKEN Genomic Sciences Center (GSC)  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: Humdraft18  
Center clone name: RP11-869L2  
----- Summary Statistics  
Sequencing vector: PCR products; 100% of reads  
Chemistry: Dye-terminator ET-amersham; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 197367 bases at least Q40  
Consensus quality: 198735 bases at least Q30  
Consensus quality: 199244 bases at least Q20  
Insert size: 199774; sum-of-contigs  
Quality coverage: 9.59x in Q20 bases; sum-of-contigs  
-----

NOTE: This is a 'working draft' sequence. It currently consists of  
11 contigs. The true order of the pieces is not known and their  
order in this sequence record is arbitrary. Gaps between the  
contigs are represented as runs N, but the exact sizes of the gaps  
are unknown. This record will be updated with the finished sequence  
as soon as it is available and the accession number will be  
preserved.

1 36889 contig of 36889 bp in length  
36990 73347 contig of 36358 bp in length  
73448 98904 contig of 25457 bp in length  
99005 124595 contig of 25591 bp in length  
124696 145130 contig of 20435 bp in length  
145231 163732 contig of 18502 bp in length  
163833 181811 contig of 17979 bp in length  
181912 190895 contig of 8984 bp in length  
190996 198298 contig of 7303 bp in length  
198399 199648 contig of 1250 bp in length  
199749 200774 contig of 1026 bp in length.

\* NOTE: This is a 'working draft' sequence.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

FEATURES  
Location/Qualifiers  
1..200774  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"









```

CDS
/replace="tct"
join(1601..1711,13695..>13736)
/gene="dJ821D11.1"
/note="other possible startcodon at 1583
this gene and dJ821D11.2 could be part of one gene
match: proteins: Tr:O95505"
/codon_start=1
/evidence=not_experimental
/product="dJ821D11.1 (PUTATIVE protein)"
/protein_id="CAAL6279.1"
/db_xref="GI:4200330"
/db_xref="SPTREMBL:O95505"
/translation="MLALTAKDSPRTALICSAWLLTASFSAQHKGSLQKDLLSQ
ACVGCLE"
variation
1760..1762
/gene="dJ821D11.1"
/note="clone CTA-109G6
gag in this entry
substitution"
1773..1923
/replace="ggg"
repeat_region
1901..1903
/note="MIR repeat: matches 47..201 of consensus"
variation
1901..1903
/gene="dJ821D11.1"
/note="clone CTA-109G6
tca in this entry
substitution"
2010..2209
/replace="tta"
repeat_region
2010..2209
/note="12 repeat: matches 2250..2453 of consensus"
variation
2302..2304
/gene="dJ821D11.1"
/note="clone CTA-109G6
gga in this entry
substitution"
3042..3152
/replace="gaa"
repeat_region
3042..3152
/note="L1M1 repeat: matches 6231..6332 of consensus"
3162..3457
/note="AluSp repeat: matches 1..297 of consensus"
3458..3551
/note="MIR repeat: matches 74..148 of consensus"
3552..3848
/note="AluDb repeat: matches 1..299 of consensus"
3849..3887
/note="MIR repeat: matches 32..74 of consensus"
4753..4754
/gene="dJ821D11.1"
/note="clone CTA-109G6
tt in this entry
deletion"
4772..4790
/replace="ttgagt"
repeat_region
4772..4790
/note="L1M4 repeat: matches 3167..3185 of consensus"
4791..5099
/note="AluSp repeat: matches 2..312 of consensus"
4791..4792
/gene="dJ821D11.1"
/note="clone CTA-109G6
tt in this entry
deletion"
5100..5116
/replace="ttttatttat"
repeat_region
5100..5116
/note="L1M4 repeat: matches 3151..3167 of consensus"
5107..5112
/gene="dJ821D11.1"
/note="clone CTA-109G6
tttat in this entry
insertion"
5117..5427
/replace="tt"
repeat_region
5117..5427
/note="AluYb8 repeat: matches 1..318 of consensus"
5359..5361

```

```

repeat_region
5430..5729
/note="AluSp repeat: matches 1..300 of consensus"
5730..5764
/note="L1M4 repeat: matches 3116..3151 of consensus"
5761..5762
/gene="dJ821D11.1"
/note="clone CTA-109G6
at in this entry
deletion"
5765..6075
/replace="att"
repeat_region
5765..6075
/note="AluY repeat: matches 1..311 of consensus"
5772..5773
/gene="dJ821D11.1"
/note="clone CTA-109G6
tt in this entry
deletion"
5778..5779
/replace="tat"
variation
5778..5779
/gene="dJ821D11.1"
/note="clone CTA-109G6
tt in this entry
deletion"
5780..5782
/replace="ttat"
variation
5780..5782
/gene="dJ821D11.1"
/note="clone CTA-109G6
tat in this entry
substitution"
5799..5801
/replace="ttt"
variation
5799..5801
/gene="dJ821D11.1"
/note="clone CTA-109G6
cgg in this entry
substitution"

```

Query Match : 55.0%; Score 22; DB 77; Length 76727;

Best Local Similarity 73.7%; Pred.No.59; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ctccagccacccataaccctcaatactccaggatg 38

Db 34580 CTTCAGACACCAACCCCTCCGACACACAGGTTTG 34543

RESULT 15

LOCUS HSAJ9611/c

DEFINITION Homo sapiens chromosome 19 clone PAC RPI-1 155110, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 3 unordered pieces.

ACCESSION AJ009611

VERSION AJ009611.5 GI:9211527

KEYWORDS HTG: HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 110387)

AUTHORS Radelof,U., Hennig,S., Ramser,J., Francis,F., Steffens,C.,

Klein,M., Seranski,P., Poustka,A., Reinhardt,R. and Lehrach,H.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 110387)

AUTHORS MPIMG.

TITLE Direct Submission

JOURNAL Submitted (17-JUL-1998) MPIMG, Abt.Lehrach, Max Planck Institut

fuer Molekulare Genetik, Ihnestrasse 73, Berlin, 14195, Germany

On Jul 15, 2000 this sequence version replaced gi:8248730.

contig 91 1..40399

```

contig 02      40500.. 61762
contig 03      61863.. 110387
Clone received from the Resource Centre of the Human Genome Project
at the Max-Planck-Institute for Molecular Genetics.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 40399: contig of 40399 bp in length
* 40400 40499: gap of 100 bp
* 40500 61762: contig of 21263 bp in length
* 61763 61862: gap of 100 bp
* 61863 110387: contig of 48525 bp in length.

FEATURES             Location/Qualifiers
     1..110387
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /chromosome="19"
         /clone="PAC RPC1-1 155110"
         /clone_lib="RPC11,3-5 Human PAC library, originating
         insitute: Roswell Park Cancer Institute, creator: Pieter
         de Jong, P. Ioannou"

misc_feature      1
                 /note="T7_end:PAC RPC1-1 155110"

misc_feature      110387
                 /note="SP6_end:PAC RPC1-1 155110"

BASE COUNT      26958 a 27876 c 28291 g 27062 t 200 others
ORIGIN

```

Query Match	55.0%	Score 22;	DB 84;	Length 110387;
Best Local Similarity	73.1%;	Pred. No. 59;		
Matches	28;	Conservative	0;	Mismatches 10; Indels 0; Gaps 0;
QY	2	tcgagccacccccaataccctcaatactccaggggattgg	39	
Db	15347	TGGAGTCATCCCATGACCCCAAGACATCCAGTGCAGG	15310	

Search completed: January 18, 2001, 04:21:48  
Job time: 10474 sec

**THIS PAGE BLANK (USPTO)**



PT cellular apoptotic a

## RESULT 1

ID V59067 standard; cDNA; 2634 BP.

AC V59067;

DT 02-FEB-1999 (first entry)

Bax inhibitor BI-1 cDNA.

KW Bax inhibitor; BI-1; human; apoptosis; ss.

OS Homo sapiens.

AA	Key	Location/Qualifiers
FH		

FT	CDS	73..786
73..786	73..786	73..786

```
FT
:/*tag= a
```

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

PN W09840397-A1.  
VV

XX  
PD  
17-SEP-1998

FD 17-SEP-1998.  
XX  
PF 13-MAR-1998; 98WO-US05015.

XX

PR 14-MAR-1997; 97US-0

[illegible]PA (BURN-) BURNHAM INST  
XXXX  
PT  
Reed, J.C. XII O:XX  
ET  
RECU  
OC, ad. V.

DR WPI; 1998-531519/45.

DR P-PSDB; W73136.

XX

PT Bax inhibitor protein

PT cellular apoptotic a



PI Amalfitano A, Chamberlain JS, Hartigan-O'Connor DJ;  
 PI Hauser MA, Kumar-Singhr;  
 XX  
 DR WPI; 1998-261485/23.

XX New adenoviral recombinant plasmid(s) - comprise sequences provided  
 PT for expression of large foreign DNA fragments, used for, e.g. gene  
 PT therapy of genetic disease(s)

XX Example 6a; Page 105-107; 139pp; English.

XX This nucleotide sequence comprises the promoter/enhancer region  
 CC of the murine muscle creatine kinase (MCK) gene. The invention  
 CC provides encapsidated adenovirus minichromosomes (EAMs) containing  
 CC a full-length murine dystrophin cDNA. The EAM consists of an  
 CC infectious encapsidated linear genome containing adenovirus  
 CC origins of replication, packaging signal elements, a  
 CC beta-galactosidase reporter gene cassette and a full-length  
 CC (14 kb) dystrophin cDNA regulated by the MCK enhancer/promoter.  
 CC EAMs are generated by cotransfecting 293 cells with supercoiled  
 CC plasmid DNA (pAd5 beta-dys) containing an embedded inverted  
 CC origin of replication (and the remaining above elements) with  
 CC linear DNA from E1-deleted virions expressing human placental  
 CC alkaline phosphatase (hpAP). All proteins necessary for the  
 CC generation of EAMs are provided in trans from the hpAP virions.  
 CC The EAMs are useful for gene transfer to a variety of cell  
 CC types in vitro and in vivo. The invention provides improved  
 CC adenovirus vectors (see also V07261) and packaging cell lines  
 CC useful for a wide variety of gene therapy applications.

XX Sequence 3364 BP; 860 A; 900 C; 810 G; 794 T; 0 other;

Query Match 50.0%; Score 20; DB 19; Length 3364;  
 Best Local Similarity 72.2%; Pred. No. 18;  
 Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 5 agccaccataaccctcaataactcaggattggg 40  
 ||||| ||| ||||| || ||||| |||  
 Db 2326 AGCCTGCCCTCACCCTGGATCCACGAGGACAGG 2291

RESULT 4  
 Q62186  
 ID Q62186 standard; DNA; 2832 BP.

XX AC Q62186;

XX DT 22-NOV-1994 (first entry)

XX DE Cyanobacteria replication protein coding sequence.

XX KW Plasmid; pAQ1; Cyanobacteria; PCC7002; vector; expression;  
 KW carbon dioxide; air; ss.

XX OS Agmenellum quadruplicatum PR-6.

XX PN JP0607877-A.

XX PD 22-MAR-1994.

XX PF 02-SEP-1992; 92JP-0274768.

XX PR 02-SEP-1992; 92JP-0274768.

XX PA (KANT ) KANSAI DENRYOKU KK.

XX DR WPI; 1994-131278/16.

DR P-PSDB; R51499.

XX New vector from cyanobacteria. - used for foreign gene expression  
 PT in cyanobacteria

XX

PS Disclosure; Page 15-18; 23pp; Japanese.

XX This sequence represents a fragment of the plasmid pAQ1 and encodes  
 CC a protein which has replication activity in Cyanobacteria. This  
 CC sequence was used in the production of a vector for the expression  
 CC of heterologous proteins in Cyanobacteria. This allows effective use  
 CC of carbon dioxide in the air.

XX Sequence 2832 BP; 760 A; 715 C; 724 G; 633 T; 0 other;

Query Match 49.5%; Score 19.8; DB 15; Length 2832;  
 Best Local Similarity 69.2%; Pred. No. 21;  
 Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 tcgagccaccataaccctcaataactcaggattggg 40  
 ||||| ||| ||||| || ||||| |||

Db 305 tcgagctaccagaacccgatcaaggcttcgggaatggg 343

RESULT 5

Q62185/C  
 ID Q62185 standard; DNA; 4809 BP.

XX AC Q62185;

XX DT 22-NOV-1994 (first entry)

XX DE pAQ1 plasmid fragment.

XX KW Plasmid; pAQ1; Cyanobacteria; PCC7002; vector; expression;  
 KW carbon dioxide; air; ss.

XX OS Agmenellum quadruplicatum PR-6.

XX PN JP06078777-A.

XX PD 22-MAR-1994.

XX PF 02-SEP-1992; 92JP-0274768.

XX PR 02-SEP-1992; 92JP-0274768.

XX PA (KANT ) KANSAI DENRYOKU KK.

XX DR WPI; 1994-131278/16.

XX New vector from cyanobacteria. - used for foreign gene expression  
 PT in cyanobacteria

XX PS Disclosure; Page 9-14; 23pp; Japanese.

XX This sequence represents a fragment of the plasmid pAQ1. This  
 CC plasmid was prepared from Cyanobacteria PCC7002. This sequence was  
 CC used in the production of a vector for the expression of heterologous  
 CC proteins in Cyanobacteria. This allows effective use of carbon  
 CC dioxide in the air.

XX Sequence 4809 BP; 1136 A; 1193 C; 1165 G; 1315 T; 0 other;

Query Match 49.5%; Score 19.8; DB 15; Length 4809;  
 Best Local Similarity 69.2%; Pred. No. 24;  
 Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 2 tcgagccaccataaccctcaataactcaggattggg 40  
 ||||| ||| ||||| || ||||| |||

Db 2348 TCGAGCTACCAAGAACCGATCAGGCTTCTGGGAATGG 2310

RESULT 6

X37371/c  
 ID X37371 standard; cDNA; 674 BP.

XX AC x37371;  
XX DT 06-JUL-1999 (first entry)  
XX DE Human secreted protein cDNA fragment containing gene 3.  
XX KW Human; secreted protein; prevention; treatment; protein therapy;  
XX KW gene therapy; diagnosis; cancer; tumour; neurodegenerative disorder;  
XX KW developmental abnormality; foetal deficiency; blood disorder; lymphoma;  
XX KW leukemia; immune system disorder; autoimmune disease; hepatic disease;  
XX KW renal disease; inflammation; allergy; asthma; sepsis; diabetes; AIDS;  
XX KW Alzheimer's disease; cognitive disorder; schizophrenia; osteoporosis;  
XX KW arthritis; psoriasis; digestive; endocrine; infection; ss.  
XX OS Homo sapiens.  
XX PN WO9909155-A1.  
XX PD 25-FEB-1999.  
XX PF 18-AUG-1998; 98WO-US17044.  
XX PR 16-JUN-1998; 98US-0092956.  
XX PR 15-JUL-1998; 98US-0092956.  
XX PR 19-AUG-1997; 97US-0056368.  
XX PR 19-AUG-1997; 97US-0056369.  
XX PR 19-AUG-1997; 97US-0056535.  
XX PR 19-AUG-1997; 97US-0056535.  
XX PR 19-AUG-1997; 97US-0056535.  
XX PR 19-AUG-1997; 97US-0056536.  
XX PR 19-AUG-1997; 97US-0056628.  
XX PR 19-AUG-1997; 97US-0056629.  
XX PR 19-AUG-1997; 97US-0056726.  
XX PR 19-AUG-1997; 97US-0056728.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Brewer LA, Duan R, Ebner R, Endress GA, Feng P;  
XX PI Florence C, Florence KA, Komatsoulis GA, Lafleur DM;  
XX PI Moore PA, Olsen HS, Rosen CA, Ruben SM, Shi Y, Soppet DR;  
XX PI Young PE;  
XX WPI; 1999-190160/16.  
XX P-PSDB; Y07746.  
XX PT New isolated human genes and the secreted polypeptides they encode  
XX PT useful for diagnosis and treatment of e.g. cancers, neurological  
XX PT disorders, immune diseases, inflammation or blood disorders  
XX PS Claim 1a; Page 192; 280pp; English.  
XX CC This invention describes novel isolated human secreted proteins and  
XX CC their encoding nucleic acid sequences. The products of the invention  
XX CC are useful for preventing, treating or ameliorating medical conditions  
XX CC e.g. by protein or gene therapy. Also pathological conditions can be  
XX CC diagnosed by determining the presence or amount of expression of  
XX CC the new polypeptides in a sample or by determining the presence or  
XX CC absence of mutations in the new polynucleotides. Specific uses are  
XX CC described for each of the 70 polynucleotides, based on which tissues they  
XX CC are most highly expressed in, and include developing products for the  
XX CC diagnosis or treatment of cancer, tumours, neurodegenerative  
XX CC disorders, developmental abnormalities and foetal deficiencies, blood  
XX CC disorders, leukemias, diseases of the immune system, autoimmune diseases,  
XX CC hepatic and renal disease, lymphomas, inflammation, allergies, asthma,  
XX CC sepsis, diabetes, Alzheimer's and cognitive disorders, schizophrenia,  
XX CC osteoporosis, arthritis, psoriasis, digestive/endocrine disorders,  
XX CC infections and AIDS. The human secreted proteins of the invention are  
XX CC represented in Y07744-Y07850 and the encoding nucleic acids are  
XX CC represented in X37369-X37441.  
XX SQ Sequence 674 BP; 238 A; 112 C; 175 G; 145 T; 4 other;

Query Match ; 48.5%; Score 19.4; DB 20; Length 674;  
Best Local Similarity 70.3%; Pred. No. 24;  
Matches 26; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
QY 4 gagccaccataaccctcaataactccaggattggg 40  
DB 645 GAGTCCTCCCACTCACCTTCAATACAAATGGCATGGG 609  
RESULT 7  
X16675/c  
ID X16675 standard; DNA; 1554 BP.  
XX AC X16675;  
XX DT 29-APR-1999 (first entry)  
XX DE Xenopus WAS45 protein encoding DNA.  
XX KW Xenopus; WAS45 protein; TGF-beta; transforming growth factor beta;  
XX KW proliferation; differentiation; mesodermal tissue; neural; muscle;  
XX KW bone; cartilage; connective tissue; wound healing; gene therapy; ss.  
XX OS Xenopus sp.  
XX FH Key Location/Qualifiers  
XX FT CDS 55..1119  
XX FT /\*tag= a  
XX FT /product= "WAS45 protein"  
XX FT sig\_peptide 55..774  
XX FT /\*tag= b  
XX FT mat\_peptide 775..1116  
XX FT /\*tag= c  
XX PN WO9902678-A1.  
XX PD 21-JAN-1999.  
XX PF 24-APR-1998; 98WO-US08334.  
XX PR 10-JUL-1997; 97US-0890918.  
XX (GEMY ) GENETICS INST INC.  
XX (WHEED ) WHITEHEAD INST BIOMEDICAL RES.  
XX PI Lavallie ER, Kacie LA, Sive H, Sun B;  
XX WPI; 1999-120879/10.  
XX P-PSDB; W94656.  
XX PT New isolated Xenopus WAS45 DNA - which is a member of the  
XX PT transforming growth factor-beta superfamily, used to induce tissue  
XX PT formation and in wound healing  
XX PS Claim 1; Page 51-52; 73pp; English.  
XX CC The present sequence encodes the Xenopus WAS45 protein. WAS45 proteins  
XX CC have the ability to induce, enhance and/or inhibit the formation,  
XX CC growth, proliferation, differentiation, maintenance of mesodermal  
XX CC tissue, including neurons and/or related neural cells and tissues such  
XX CC as brain cells, Schwann cells, glial cells and astrocytes, as well as  
XX CC muscle cells and tissues. They can be used for treating bone, cartilage,  
XX CC muscle, nerve, epidermis and/or other connective tissue defects, as well  
XX CC as periodontal disease and healing of various epidermis, nerve, smooth  
XX CC including spinal cord, muscle, including cardiac, striated or smooth  
XX CC muscle, and other tissues and wounds, and other organs such as liver,  
XX CC pancreas, spleen brain, lung, cardiac and kidney tissue. They can also  
XX CC be used to treat or prevent such conditions as osteoarthritis,  
XX CC osteoporosis, and other abnormalities of bone, cartilage, muscle, nerve,  
XX CC epidermis or other connective tissue, organs such as liver, pancreas,  
XX CC spleen, lung, cardiac and kidney and other tissues. They can also be  
XX CC used for wound healing, reduction of fibrosis and reduction of scar  
XX CC tissue formation. They can also be used to induce bone and/or cartilage

The present invention describes a library of human polynucleotides comprising the sequences given in 212532 to 217779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in 212532 to 217779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels,



```
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 1 ctcgagccaccataaccctcaataactccaggattggg 40
   ||| ||| |||| | ||| |||| | ||||
Db 955 ctcgagcattccaaagctcttaattctccataaaatggg 994

RESULT 12
X86274/c
ID X86274 standard; DNA; 2302 BP.
XX
AC X86274;
XX
DT 22-SEP-1999 (first entry)
XX
DE DNA encoding a human PIG protein.
XX
KW p53 transcription tag; p53 status; cancer; cytotoxicity; carcinogenicity;
KW neoplastic; PIG; ss.
XX
OS Homo sapiens.
XX
PN W09914356-A2.
XX
PD 25-MAR-1999.
XX
PF 17-SEP-1998; 98WO-US19300.
XX
PR 30-MAR-1998; 98US-0079817.
PR 17-SEP-1997; 97US-0059153.
XX
PA (UYJO ) UNIV JOHNS HOPKINS.
XX
PI Kinzler KW, Polyak K, Vogelstein B;
XX
DR WPI; 1999-443793/37.
XX
PT Use of p53 transcription tags to determine p53 status in, e.g.
PT cancer diagnosis
XX
PS Disclosure; Page 65-66; 73pp; English.
XX
CC The specification describes the use of p53 transcription tags for
CC developing products to determine p53 status, to diagnose cancer
CC and to evaluate cytotoxicity or carcinogenicity of a test agent.
CC A method for diagnosing cancer or determining p53 status in a sample
CC suspected for being neoplastic comprises comparing the level of
CC transcription of an RNA transcript in a first sample (s1) of a first
CC tissue (t1) to the level of transcription of the transcript in a second
CC sample (s2) of a second tissue (s2), where s1 is suspected of being
CC neoplastic and s2 is a normal human tissue (of the same type) and the
CC transcript is identified by a tag; and categorizing s1 as neoplastic
CC or as having a mutant p53 when transcription is found to be the same
CC or lower in the first, than in s2. The methods and products can be used
CC to determine p53 status, to diagnose cancer and to evaluate cytotoxicity
CC or carcinogenicity of a test agent. X86265-75 encode human PIG proteins.
XX
SQ Sequence 2302 BP; 497 A; 752 C; 594 G; 459 T; 0 other;
```

```
AC Q50146;
XX
DT 03-JUN-1994 (first entry)
XX
DE Restriction enzyme.
XX
KW Restriction site; enzyme; SphI; BamHI; cleavage-site;
KW phage; Escherichia coli; ss.
XX
FH Key :Location/Qualifiers
FT mat_peptide 240..1074 /*tag= b
FT misc_feature 897 /*tag= c
FT /*label= SphI_cleavage-site
FT mat_peptide 1127..2008 /*tag= d
FT misc_feature 1425 /*tag= e
FT /*label= BamHI_cleavage-site
XX
PN JPO5244946-A.
XX
PD 24-SEP-1993.
XX
PF 03-MAR-1992; 92JP-0094990.
XX
PR 03-MAR-1992; 92JP-0094990.
XX
PA (KAJI/) KAJI A.
XX
DR WPI; 1993-389087/49.
DR P-PSDB; R42305.
XX
PT New restriction enzyme - suppresses melting by T-even phage of
PT Escherichia coli
XX
PS Claim 1; Page 7-10; 10pp; Japanese.
XX
CC The enzyme can suppress the melting by T-even phage of E.coli, in a
CC temp. dependent manner. A large amt. of enzyme can be produced.
XX
SQ Sequence 2401 BP; 681 A; 564 C; 543 G; 613 T; 0 other;

Query Match 47.5%; Score 19; DB 14; Length 2401;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 3 ctagccaccataaccctcaataactccaggatt 37
   ||| |||| | ||| | |||| | ||||
Db 284 caaggcaccgcacacacgcaatacacacaggcgatt 318

RESULT 14
X14116
ID X14116 standard; DNA; 1183 BP.
XX
AC X14116;
XX
DT 31-MAR-1999 (first entry)
XX
DE H. pylori GHPO 106 gene.
XX
KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW peptic ulcer disease; ss.
XX
OS Helicobacter pylori.
XX
FH Key :Location/Qualifiers
FT CDS 51..1133 /*tag= a
XX
```

```
PN WO9843478-A1.
XX
PD 08-OCT-1998.
XX
PF 01-APR-1998; 98WO-US06371.
XX
PR 29-JUL-1997; 97US-0902615.
PR 01-APR-1997; 97US-0833457.
PR 24-JUN-1997; 97US-0881227.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
XX Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
XX WPI: 1998-542293/46.
DR P-PSDB; W98397.
XX
XX New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
XX
XX Claim 1; Page 650-651; 2054pp; English.
XX
XX This sequence represents a polynucleotide of the invention. It was
CC isolated from Helicobacter pylori and encodes a H.pylori GHPD protein.
CC The polypeptides can be used for preventing or treating Helicobacter
CC infections, and gastroduodenal diseases associated with these
CC infections, including acute, chronic, and atrophic gastritis, and peptic
CC ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX
XX Sequence 1183 BP; 413 A; 200 C; 242 G; 328 T; 0 other;

Query Match 47.0%; Score 18.8; DB 19; Length 1183;
Best Local Similarity 68.4%; Pred. No. 49;
Matches 26; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 2 tcgagccaccataaccctcaataactccaggattgg 39
Db ||||| | ||||| ||||| ||||| |||||
355 tcgtgagcgctcaataaccctcaataactcctatgaatgg 392

RESULT 15
292219/C
ID 292219 standard; DNA; 2191 BP.
XX
AC 292219;
XX
XX 05-JUN-2000 (first entry)
XX
XX Human thioredoxin reductase nucleotide sequence #1.
XX
XX Thioredoxin reductase; redox inhibitor; cellular proliferation; cancer;
KW apoptosis; reperfusion injury; hepatitis; amyotrophic lateral sclerosis;
KW Alzheimer's disease; autoimmune disease; Sjogren's syndrome; lupus;
KW rheumatoid arthritis; HIV; Hermansky-Pudlak syndrome; retinopathy;
KW wound healing; Crohn's disease; ulcerative colitis; angiogenesis;
KW adult respiratory distress syndrome; ARDS; infection; ds;
KW inflammatory conditions; psoriasis; inflammatory bowel disease;
KW hyperproliferative disease; mycosis fungoides.
XX
OS Homo sapiens.
XX
XX WO200006088-A2.
XX
XX 10-FEB-2000.
XX
XX 02-AUG-1999; 99WO-US17496.
XX
XX 31-JUL-1998; 98US-0127219.
XX

(KIRK/) KIRKPATRICK D L.
PA (POWI/) POWIS G.
XX
XX Kirkpatrick DL, Powis G;
XX WPI: 2000-195158/17.
XX
XX New composition for preventing inhibition of apoptosis and for treating
PT e.g. cancer or reperfusion injuries comprises inhibitor of cellular
PT redox signalling
XX
XX Disclosure; Page 51-53; 55pp; English.
XX
XX This sequence represents a thioredoxin reductase nucleotide sequence.
CC Thioredoxin exerts specific redox control over a number of transcription
CC factors to modulate their DNA binding and therefore regulates gene
CC transcription. Thioredoxin reductase catalyses the reduction of
CC thioredoxin. The thioredoxin redox system plays an important role in a
CC number of diseased states. The invention relates to a composition
CC comprising an inhibitor of cellular redox signalling and a carrier. The
CC composition can be used for inhibiting cellular proliferation or
CC preventing inhibition of apoptosis and is therefore useful for treating
CC cancer, reperfusion injury following ischaemia, hepatitis, amyotrophic
CC lateral sclerosis, neurodegenerative diseases, Alzheimer's disease,
CC autoimmune diseases, Sjogren's syndrome, lupus, rheumatoid arthritis,
CC HIV, Hermansky-Pudlak syndrome, retinal oxidative damage, retinopathy,
CC skin hyperplasia, aging, ultraviolet damage, wound healing, Crohn's
CC disease, ulcerative colitis, angiogenesis, uterine disorders, adult
CC respiratory distress syndrome (ARDS), lung disorders, viral infections
CC such as herpes virus, pox virus and adenovirus infections, inflammatory
CC conditions, autoimmune diseases (such as systemic lupus erythematosus,
CC rheumatoid arthritis, psoriasis, inflammatory bowel disease and
CC autoimmune diabetes), immune mediated glomerular nephritis,
CC hyperproliferative diseases (such as fibrosis) and mycosis fungoides.
XX
XX Sequence 2191 BP; 614 A; 411 C; 555 G; 611 T; 0 other;

Query Match 47.0%; Score 18.8; DB 21; Length 2191;
Best Local Similarity 76.7%; Pred. No. 55;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 5 agccacccataaccctcaataactccagg 34
Db ||||| | ||||| ||||| |||||
878 AGCCAGCCAGAACCCCTGCACACTCCAGG 849

Search completed: January 18, 2001, 11:04:27
Job time: 29436 sec
```



pct-us00-32259-12.rng

Thu Jan 18 12:41:10 2001

**THIS PAGE BLANK (USPTO)**







; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/735,609  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: UM-02484  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3364 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-09-315-372-9

Query Match 50.0%; Score 20; DB 3; Length 3364;  
Best Local Similarity 72.2%; Pred. No. 11;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 agccacccataaccctcaatactccaggattggg 40  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2326 AGCGTCCCTCACCCCTGGATCCACGAGGACAGG 2291

RESULT 8  
US-09-244-752-9/c  
; Sequence 9, Application US/09244752  
; Patent No. 6063622  
; GENERAL INFORMATION:  
; APPLICANT: Chamberlain, Jeffrey S.  
; APPLICANT: Amalfitano, Andrea  
; APPLICANT: Hauser, Michael A.  
; APPLICANT: Kumar-Singh, Rajendra  
; APPLICANT: Hartigan-O'Connor, Dennis J.  
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/244,752  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/735,609  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: UM-02484  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3364 base pairs

; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-09-244-752-9

Query Match 50.0%; Score 20; DB 3; Length 3364;  
Best Local Similarity 72.2%; Pred. No. 11;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 agccacccataaccctcaatactccaggattggg 40  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2326 AGCGTCCCTCACCCCTGGATCCACGAGGACAGG 2291

RESULT 9  
US-09-245-497-9/c  
; Sequence 9, Application US/09245497  
; Patent No. 6083750  
; GENERAL INFORMATION:  
; APPLICANT: Chamberlain, Jeffrey S.  
; APPLICANT: Amalfitano, Andrea  
; APPLICANT: Hauser, Michael A.  
; APPLICANT: Kumar-Singh, Rajendra  
; APPLICANT: Hartigan-O'Connor, Dennis J.  
; TITLE OF INVENTION: IMPROVED ADENOVIRUS VECTORS  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States Of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/245,497  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/735,609  
; APPLICATION NUMBER: <B> FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: UM-02484  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3364 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: other nucleic acid  
; DESCRIPTION: /desc = "DNA"  
US-09-245-497-9

Query Match 50.0%; Score 20; DB 3; Length 3364;  
Best Local Similarity 72.2%; Pred. No. 11;  
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 5 agccacccataaccctcaatactccaggattggg 40  
||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 2326 AGCGTCCCTCACCCCTGGATCCACGAGGACAGG 2291



; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1864  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (275)..(754)  
US-09-221-416-10

Query Match 48.0%; Score 19.2; DB 3; Length 1864;  
Best Local Similarity 67.5%; Pred. No. 21;  
Matches 27; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 ctgagccacccccaaccctcaactactccaggatggg 40  
||| ||| |||| | ||| ||||| | |||  
Db 955 ctccagcattcccaagctcttaattctccataaattggg 994

RESULT 15

US-08-770-379-20  
; Sequence 20, Application US/08770379  
; Patent No. 5849564  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/770,379  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 52342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32207 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 47.0%; Score 18.8; DB 2; Length 32207;  
Best Local Similarity 76.7%; Pred. No. 54;  
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 accccataaccctcaactactccaggattg 38  
||| || ||| || | ||||| |||

Db 3464 ACCATATCACCATCCACACTCCAGGGACTG 3493

Search completed: January 18, 2001, 10:59:56  
Job time: 30474 sec





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: January 18, 2001, 03:53:45 ; Search time 3453.92 Seconds  
(without alignments)  
81.154 Million cell updates/sec

Title: PCT-US00-32259-12  
Perfect score: 40  
Sequence: 1 ctcgagccaccctcaataccctcaactccaggattggg 40

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues  
Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_estl1:\*  
2: gb\_estl2:\*  
3: gb\_estl3:\*  
4: gb\_estl4:\*  
5: gb\_estl5:\*  
6: gb\_estl6:\*  
7: gb\_estl7:\*  
8: gb\_estl8:\*  
9: gb\_estl9:\*  
10: gb\_estl10:\*  
11: gb\_estl11:\*  
12: gb\_estl12:\*  
13: gb\_estl13:\*  
14: gb\_estl14:\*  
15: gb\_estl15:\*  
16: gb\_estl16:\*  
17: gb\_estl17:\*  
18: gb\_estl18:\*  
19: gb\_estl19:\*  
20: gb\_estl20:\*  
21: gb\_estl21:\*  
22: gb\_estl22:\*  
23: gb\_estl23:\*  
24: gb\_estl24:\*  
25: gb\_estl25:\*  
26: gb\_estl26:\*  
27: gb\_estl27:\*  
28: gb\_estl28:\*  
29: gb\_estl29:\*  
30: gb\_estl30:\*  
31: gb\_estl31:\*  
32: gb\_estl32:\*  
33: gb\_estl33:\*  
34: gb\_estl34:\*  
35: gb\_estl35:\*  
36: gb\_estl36:\*  
37: gb\_estl37:\*  
38: gb\_estl38:\*  
39: gb\_estl39:\*  
40: gb\_estl40:\*  
41: em\_estba:\*  
42: em\_estfun:\*  
43: em\_esthum1:\*  
44: em\_esthum2:\*  
45: em\_esthum3:\*  
46: em\_esthum4:\*  
47: em\_esthum5:\*  
48: em\_esthum6:\*  
49: em\_esthum7:\*  
50: em\_esthum8:\*  
51: em\_esthum9:\*  
52: em\_esthum10:\*  
53: em\_esthum11:\*  
54: em\_esthum12:\*  
55: em\_esthum13:\*  
56: em\_esthum14:\*  
57: em\_esthum15:\*  
58: em\_esthum16:\*  
59: em\_esthum17:\*  
60: em\_esthum18:\*  
61: em\_esthum19:\*  
62: em\_esthum20:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estov1:\*  
68: em\_estov2:\*  
69: em\_estpl1:\*  
70: em\_estpl2:\*  
71: em\_estpl3:\*  
72: em\_estpl4:\*  
73: em\_estpl5:\*  
74: em\_estro1:\*  
75: em\_estro2:\*  
76: em\_estro3:\*  
77: em\_estro4:\*  
78: em\_estro5:\*  
79: em\_estro6:\*  
80: em\_estro7:\*  
81: em\_estro8:\*  
82: em\_estro9:\*  
83: em\_estro10:\*  
84: em\_estro11:\*  
85: em\_estro12:\*  
86: em\_estro13:\*  
87: gb\_est41:\*  
88: gb\_est42:\*  
89: gb\_est43:\*  
90: gb\_est44:\*  
91: gb\_est45:\*  
92: gb\_est46:\*  
93: gb\_est47:\*  
94: gb\_est48:\*  
95: gb\_est49:\*  
96: gb\_est50:\*  
97: gb\_est51:\*  
98: gb\_est52:\*  
99: gb\_est53:\*  
100: gb\_est54:\*  
101: gb\_est55:\*  
102: gb\_est56:\*  
103: gb\_est57:\*  
104: gb\_est67:\*  
105: gb\_est68:\*  
106: gb\_est69:\*  
107: gb\_est70:\*  
108: gb\_est71:\*  
109: gb\_est72:\*  
110: gb\_est73:\*  
111: gb\_est74:\*  
112: em\_esthum21:\*  
113: em\_esthum22:\*  
114: em\_esthum23:\*  
115: em\_estom1:\*  
116: em\_estom2:\*

117: em\_estpl6:\*  
 118: em\_estpl7:\*  
 119: em\_estpl8:\*  
 120: em\_estro14:\*  
 121: em\_estro15:\*  
 122: em\_estro16:\*  
 123: em\_estro17:\*  
 124: em\_estro18:\*  
 125: em\_estro19:\*  
 126: gb\_est58:\*  
 127: gb\_est59:\*  
 128: gb\_est60:\*  
 129: gb\_est61:\*  
 130: gb\_est62:\*  
 131: gb\_est63:\*  
 132: gb\_est64:\*  
 133: gb\_est65:\*  
 134: gb\_est66:\*  
 135: gb\_est75:\*  
 136: gb\_est76:\*  
 137: gb\_est77:\*  
 138: gb\_est78:\*  
 139: gb\_est79:\*  
 140: gb\_est80:\*  
 141: gb\_est81:\*  
 142: gb\_est82:\*  
 143: gb\_est83:\*  
 144: gb\_est84:\*  
 145: gb\_est85:\*  
 146: gb\_est86:\*  
 147: gb\_est87:\*  
 148: gb\_gss1:\*  
 149: gb\_gss2:\*  
 150: gb\_gss3:\*  
 151: gb\_gss4:\*  
 152: em\_gss1:\*  
 153: em\_gss2:\*  
 154: em\_gss3:\*  
 155: em\_gss4:\*  
 156: gb\_gss5:\*  
 157: gb\_gss6:\*  
 158: gb\_gss7:\*  
 159: gb\_gss8:\*  
 160: gb\_gss9:\*  
 161: em\_gss5:\*  
 162: em\_gss6:\*  
 163: em\_gss7:\*  
 164: em\_gss8:\*  
 165: em\_gss9:\*  
 166: em\_gss10:\*  
 167: em\_gss11:\*  
 168: gb\_gss10:\*  
 169: gb\_gss11:\*  
 170: em\_gss12:\*  
 171: gb\_gss12:\*  
 172: gb\_gss13:\*  
 173: gb\_gss14:\*  
 174: gb\_gss15:\*  
 175: gb\_gss16:\*  
 176: gb\_gss17:\*  
 177: gb\_gss18:\*  
 178: gb\_gss19:\*  
 179: em\_gss13:\*  
 180: gb\_gss20:\*  
 181: gb\_gss21:\*  
 182: gb\_gss22:\*  
 183: gb\_gss23:\*  
 184: gb\_gss24:\*  
 185: em\_gss14:\*  
 186: em\_gss15:\*  
 187: em\_gss16:\*  
 188: em\_gss17:\*  
 189: em\_gss18:\*

190: gb\_gss25:\*  
 191: gb\_gss26:\*  
 192: gb\_gss27:\*  
 193: gb\_gss28:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24.6	61.5	296	96 AW845395	AW845395 CM4-CT004
C 2	23.8	59.5	250	145 T08487	T08487 EST06378 In
C 3	23.6	59.0	337	146 T92497	T92497 Y024009.r1
C 4	23	57.5	69	137 BE878784	BE878784 601493062
C 5	23	57.5	118	95 AW804230	AW804230 PM3-UM008
C 6	23	57.5	135	133 BE005490	BE005490 CM1-BN011
C 7	23	57.5	147	133 BE005492	BE005492 CM1-BN011
C 8	23	57.5	161	143 N73434	N73434 Y231e12.r1
C 9	23	57.5	223	134 BE089733	BE089733 QV0-BT070
C 10	23	57.5	227	133 BE005486	BE005486 CM1-BN011
C 11	23	57.5	231	5 AA327421	AA327421 EST30756
C 12	23	57.5	237	139 BF037734	BF037734 601462081
C 13	23	57.5	249	95 AW804233	AW804233 PM3-UM008
C 14	23	57.5	251	110 BE622136	BE622136 601440862
C 15	23	57.5	252	136 BE869246	BE869246 601445061
C 16	23	57.5	252	137 BE894967	BE894967 601436172
C 17	23	57.5	260	2 AA101534	AA101534 zn79c11.s
C 18	23	57.5	263	6 AA383010	AA383010 EST96488
C 19	23	57.5	265	137 BE878754	BE878754 601493017
C 20	23	57.5	276	110 BE617770	BE617770 601441968
C 21	23	57.5	289	145 T39447	T39447 Y066508.r1
C 22	23	57.5	299	104 BE178852	BE178852 PM1-HT060
C 23	23	57.5	304	6 AA384401	AA384401 EST97914
C 24	23	57.5	305	12 AA853875	AA853875 NHT8Cae09
C 25	23	57.5	307	95 AW804241	AW804241 PM3-UM008
C 26	23	57.5	307	145 T39458	T39458 Y066507.r2
C 27	23	57.5	308	6 AA383479	AA383479 EST96859
C 28	23	57.5	308	95 AW804238	AW804238 PM3-UM008
C 29	23	57.5	313	137 BE936437	BE936437 RC1-NT003
C 30	23	57.5	326	6 AA361993	AA361993 EST71440
C 31	23	57.5	328	95 AW804271	AW804271 PM3-UM008
C 32	23	57.5	329	135 BE763332	BE763332 RC1-NT003
C 33	23	57.5	332	6 AA373764	AA373764 EST85818
C 34	23	57.5	334	144 R66220	R66220 Y134807.r1
C 35	23	57.5	335	145 T52898	T52898 Y081112.r1
C 36	23	57.5	337	38 AV703130	AV703130 AV703130
C 37	23	57.5	342	6 AA363971	AA363971 EST74461
C 38	23	57.5	343	137 BE936435	BE936435 RC1-NT003
C 39	23	57.5	344	95 AW804222	AW804222 PM3-UM008
C 40	23	57.5	348	134 BE089684	BE089684 QV0-BT070
C 41	23	57.5	348	134 BE089688	BE089688 QV0-BT070
C 42	23	57.5	349	134 BE089689	BE089689 QV0-BT070
C 43	23	57.5	353	133 BE003833	BE003833 QV3-BN009
C 44	23	57.5	354	96 AW904630	AW904630 RC1-NN106
C 45	23	57.5	361	133 BE005530	BE005530 CM1-BN011

## ALIGNMENTS

RESULT 1  
 AW845395/c

LOCUS CM4-CT0045-180200-512-a09 CT0045 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION AW845395  
 ACCESSION AW845395  
 VERSION AW845395.1 GI:7940912  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens



PM3-UM00089-170300-002-d06 UM0089 Homo sapiens CDNA, mRNA sequence.  
AW804230. GI:7856100  
EST.  
human.

1. (Bases 1 to 118)  
Mammalia, Eutheria, Primates, Cebidae, Hominidae, Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. J., Zago, M. A., Bordin, S., Costa, F. F., Carvalho, G. H., Goldman, A. F., Matsukuma, A., Baia, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Buchner, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Received: 1999  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202863  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research

Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: [zslimponse@ludwig.org.br](mailto:zslimponse@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml.pl?tbl=ets3-PM3-UM0089-17>)  
300-02-4066t3-2000-03-17t4-1  
Seq primer: puc 18 forward  
High quality sequence start: 23  
High quality sequence stop: 118.

```

1. .118
;/organism="Homo sapiens"
/db_xref="taxon:9606"
;/clone_lib="UK0089"
/dev_stage="Adult"
/note="Organ: uterus; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 19/716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue

```

mRNA and cDNA amplification were performed under low stringency conditions.

```

25 a      31 c      35 g      27 t
          57.5%; Score 23; DB 95; Length 118;
Similarity 74.4%; Pred. No. 36;
29; Conservative 0; Mismatches 10; Indels 0; Gaps 0
ggagccacccataaacctcaatactccaggattggg 40
|||||

```

BE005490: 135 bp mRNA EST 05-JUN-2000  
CM1-BN0117-110400-183-g08 BN0117 Homo sapiens CDNA, mRNA sequence.

BE003490  
BE005490:1 GI:8265723  
EST.  
human.

1 (bases 1 to 135)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R.,  
Eukaryota: Metazoa;  
Mammalia: Eutheria;  
Primates; Catarrhini; Homnidae: Homo.











Search completed: January 18, 2001, 03:53:47  
Job time: 8901 sec

**THIS PAGE BLANK (USPTO)**